UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/774,925	01/31/2001	Sara H. Basson	YOR920000740US1	5320

7590 02/27/2004

Ryan, Mason & Lewis, LLP
90 Forest Avenue
Locust Valley, NY 11560

EXAMINER	
BRANT, DMITRY	

ART UNIT	PAPER NUMBER
2655	7

DATE MAILED: 02/27/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/774,925	BASSON ET AL.	
	Examiner	Art Unit	
	Dmitry Brant	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *01/31/2001*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-28* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-28* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *4,5*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. The U.S. patents of Butnaru et al., Potts et al. and Van Schyndel teach computer-based apparatuses (systems) and hence the methods and computer code necessary to implement these systems are inevitably part of their teachings.

3. Claims 1-2, 6-8, 10-17, 19-22 are rejected under 35 U.S.C. 103(a) as being unpatentable over Butnaru et al. (6,240,392) in view of Potts et al. (6,593,956), and further in view of Jhabvala et al. (5,029,216)

As per claim 1 and 15, Butnaru et al. disclose a wearable device for people with hearing disabilities. The system comprises a wearable computer that is capable of displaying a variety of indicators (elem. 80, 90, FIG. 2)

Butnaru et al. do not disclose:

- identifying the location of the individual who is currently speaking during the event;
- determining whether the individual identified as the current speaker is within a field of view of the user;

- displaying a first visual indicator to the user, in accordance with the display system, in association with the individual identified as the current speaker when the individual is within the field of view of the user;
- displaying a second visual indicator to the user, in accordance with the display system, when the individual identified as the current speaker is not within the field of view of the user.

Potts et al. teach identifying the speaking individual using audio signals and video images and, if the speaker is not in the view, changing the camera view to his/her location (Col. 6, line 26-34). Potts et al. also teach the use of video camera (elem. 14, FIG. 3) for capturing images of people participating in the conference and a video based locator (elem. 60, FIG. 3) for processing the images.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Butnaru et al. as taught by Potts et al. in order create a wearable device containing a camera capable of locating, identifying and potentially zooming to the current speaker. This would allow the wearer of the system to identify and observe the current speaker through the display (viewfinder) of the wearable device.

Potts et al. do not teach

- displaying a first visual indicator to the user, in accordance with the display system, in association with the individual identified as the current speaker when the individual is within the field of view of the user;
- displaying a second visual indicator to the user, in accordance with the display system, when the individual identified as the current speaker is not within the field of view of the user.

Jhabvala et al. teach the use of visual indicators to instruct the deaf users about the direction of incoming sounds (Col. 3, lines 2-8).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify a wearable device of Butnaru, modified with a speaker-locating capability taught by Potts et al. to utilize visual indicators as taught by Jhabvala et al., in order to allow the wearable system to identify the direction of incoming speech, and if the speaker was not in view, using visual indicators to instruct the user to turn his/her head in the appropriate direction. Similar to the many well-known uses of various indicators in normal life (turn signals in cars, green/red arrows on the intersection lights, highlighting of selected items on Windows, etc), the use of visual indicators with the display would assist deaf people in determining the direction of incoming speech and focusing on the current speaker.

As per claims 2, 16, and 17, Butnaru et al. disclose a head-mounted display system that is wearable by the user (elem. 120, FIG.1)

As per claim 6, Butnaru et al. do not disclose the step of identifying location of the individual based on audio information.

Potts et al. teach the use of audio-based locator (elem. 70, FIG. 3) that captures and processes audio signals from a microphone array and then determines the location of the audio source using speaker location module (elem. 114, FIG. 4 and Col. 17, lines 35-55)

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Butnaru et al. as taught by Potts et al. in order to determine the location of the current speaker using audio information. Because the current speaker is most readily identified by having generated acoustic signals, audio-based locator would be the natural (and cheapest) choice for identifying the source of the incoming speech.

As per claim 7, Butnaru et al. do not disclose "the step of determining whether the individual identified as the current speaker is within the field of view of the user further comprises capturing directional data associated with the display system and positional data associated with the user."

Potts et al. teach determining whether the current speaker is within the view of the camera by processing the results of audio speaker location module (FIG. 13 and Col. 18, lines 34-39). Because the camera is mounted on the head of the user and faces in the same direction as the user, the apparatus described by Pott et al. would unavoidably have to take the position of the user and the direction of user's view as its frame of reference.

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Butnaru et al. as taught by Potts et al. in order to determine whether the current speaker is within the field of view of the user using audio data because this would allow the system to either pinpoint the speaker or direct the user to look in some other direction in order to place the speaker within the user's view.

As per claim 8, Butnaru et al. disclose a system capable of displaying a variety of visual indicators, such as user's location (elem. 90, FIG.2)

Butnaru et al. do not disclose "displaying the first visual indicator further comprises correlating the location of the current speaker with the directional data associated with the display system and the positional data associated with the user."

Potts et al. teach determining whether the current speaker is within the view of the camera by processing the results of audio speaker location module (FIG. 13 and Col. 18, lines 34-39) by correlating the position of the speaker with the position of the microphones. Because the camera is mounted on the head of the user and faces in the same direction as the user (hence, the same location for camera, microphone and the user), the apparatus described by Pott et al. would unavoidably have to correlate the position of the user and the direction of user's view with the estimated position of the speaker.

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify visual indicators described by Butnaru et al. to notify the user of current speaker's location, based on the information supplied by the method taught by Potts et al. This would allow the user to identify the current speaker who is in his field-of-view using the information obtained from the visual indicator and hence focus his/attention on the speaker.

As per claims 10-11, 19-20, Butnaru et al. disclose a system capable of displaying a variety of visual indicators on the computer screen (FIG.2)

Butnaru et al. do not disclose "visual indicator [that] comprises a change in at least one attribute associated with a representation of the individual identified as the current speaker on the display system", where the attribute is one of color and brightness.

The examiner takes official notice that computer screens have an inherent ability to display various images of varying colors and brightness. In addition, the technique of highlighting objects on the screen to bring them into the focus is well-known to the practitioners in computer arts. For example, Windows operating system changes the color and brightness of file folder images when user selects them.

It would have been obvious to one of ordinary skill in the art at the time the invention was made that display system disclosed by Butnaru et al. to change the color and brightness of the current speaker's image in order to pinpoint the speaker to the user, because the change in the visual representation of the current speaker would quickly get attention of the user and allow him to focus on the speaker.

As per claims 12 and 21, Butnaru et al. disclose a system capable of displaying a variety of visual indicators on the computer screen (80, 90 FIG.2).

Butnaru et al. do not disclose that a "second visual indicator comprises a directional symbol displayed on the display system indicating to the user the direction to turn such that the current speaker is in the user's field of view."

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify visual indicators described by Butnaru et al. to notify the user of the current speaker's location. This would allow the user to determine whether the current speaker is in his/her field of view and, if not, turn the head in the direction shown

by the visual indicator, so as to see the speaker. Similar to the many well-known uses of various indicators in normal life (turn signals in cars, green/red arrows on the intersection lights, etc), the use of visual indicators with the display would assist the user in determining the direction of incoming speech and focusing on the current speaker.

As per claims 13-14 and 22, Butnaru et al. discloses the system that is capable of recognizing the speech using speech recognizer (elem. 55, FIG. 3) and displaying the text content to the user via a display system (Col. 5, lines 49-57).

4. Claims 3-4 are rejected under 35 U.S.C. 103(a) as being unpatentable over Butnaru et al., Potts et al. and Jhabvala et al. as applied to claim 1, and further in view of Van Schyndel (5,940,118)

As per claim 3, Potts et al. discloses a video system that captures an image of the speaker (Col. 2, lines 18-20), determining the person is currently speaking and finding his location (Col. 2, line 39). Potts's system uses a scheme that requires less processing and identifies the active speaker based on the difference of flesh tones found between the current and previous video frames.

Butnaru et al., Potts et al. and Jhabvala et al. do not disclose "analyzing the one or more capture video images to determine which individual has one or more facial feature indicative of speech."

Van Schyndel teaches capturing images of the users and identifying current speaker based on his/her head and mouth movements and determining the location of that person (Col. 7, lines 35-46).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the wearable system of Butnaru et al., Potts et al. and Jhabvala et al. (specifically the part taught by Potts et al.) as taught by Van Schyndel, in order to improve the identification of the speaker, because detection of mouth movements would produce more reliable results at the expense of heavier image processing.

As per claim 4, Butnaru et al. do not disclose that "the step of determining whether the individual identified as the current speaker is within the field of view of the user further comprises capturing one or more video images of the field of view of the user."

Potts et al. discloses capturing and adjusting the field-of-view of the speaker in order to follow the speaker (Col. 2, line 63- Col.3, line 2)

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Butnaru et al. as taught by Potts et al. in order to determine whether the individual currently speaking is within the field-of-view of the user wearing the system, because this determination would allow the system to either pinpoint the current speaker or direct the user to look in some other direction in order to place the speaker within the user's view.

5. Claim 5 is rejected under 35 U.S.C. 103(a) as being unpatentable over Butnaru et al., Potts et al. and Jhabvala et al., as applied to claim 4, and further in view of Hein et al. (6,466,250)

Butnaru et al. disclose a system capable of displaying a variety of visual indicators on the computer screen (FIG.2). Hence, the examiner takes official notice that computer screens have an inherent ability to display various images.

Potts et al. discloses capturing the image of the speaker (Col. 2, line 63- Col.3, line 2).

Butnaru et al., Potts et al. and Jhabvala et al. do not disclose that "the step of displaying the first visual indicator further comprises correlating at least a portion of the one or more video images captured of the individuals participating in the event with at least a portion of the one or more video images captured of the field of view of the user."

Hein et al. teaches placing the moving image of a speaker into the view of the user (Col. 6, lines 47-50)

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Butnaru et al., Potts et al. and Jhabvala et al. as taught by Hein et al. in order to bring the speaker to the attention of the user by identifying him using visual emphasis. This would allow the user to quickly pinpoint the active speaker without having to look for other, less obvious indicators.

6. Claims 9 and 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Butnaru et al., Potts et al., and Jhabvala et al., as applied to claim 1, and further in view of Hein et al. (6,466,250)

Butnaru et al., Pott et al., and Jhabvala et al. do not disclose that a "first visual indicator comprises a marker displayed in proximity to a representation of the individual identified as the current speaker on the display system."

Hein et al. teaches using a colored frame around the speaker's image in order to identify him to the viewer (Col. 7, lines 9-11)

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify visual indicators described by Butnaru et al. as taught by Hein et al. in order to identify the current speaker located in the field of view of the user, because the change in the visual representation of the current speaker would quickly get attention of the user and allow him to focus on the speaker.

7. Claims 23, 25-28 are rejected under 35 U.S.C. 103(a) as being unpatentable over Butnaru et al. in view of Potts et al. and further in view of Van Schyndel and Jhabvala et al.

Butnaru et al. disclose a wearable device for people with hearing disabilities. The system comprises a wearable computer that is capable of displaying a variety of indicators (elem. 80, 90, FIG. 2).

It would have been obvious to one of ordinary skill in the art at the time the invention was made that a video camera would be necessary in order to establish the user's field of view. Because the camera would be situated on the user's head, it would only capture the information located within the user's view and thus give good indication of where the user is currently looking.

Butnaru et al. do not disclose:

- one or more video cameras for capturing video images of the one or more individuals participating in the event;
- a video server coupled to the one or more video cameras and operative to: (i) analyze the captured video images to determine which individual has one or more facial features indicative of speech; and (ii) identify the location of the individual who is currently speaking during the event;
- identifying the location of the individual who is currently speaking during the event;
- determining whether the individual identified as the current speaker is within a field of view of the user;
- displaying a first visual indicator to the user, in accordance with the display system, in association with the individual identified as the current speaker when the individual is within the field of view of the user;
- displaying a second visual indicator to the user, in accordance with the display system, when the individual identified as the current speaker is not within the field of view of the user.

Potts et al. teach identifying the speaking individual using audio signals and video images and, if the speaker is not in the view, changing the camera view to his/her location (Col. 6, line 26-34). Potts et al. also teach the use of video camera (elem. 14, FIG. 3) for capturing images of people participating in the conference and a video based locator (elem. 60, FIG. 3) for processing the images.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Butnaru et al. as taught by Potts et al. in order create a wearable device containing a camera capable of locating, identifying and potentially zooming to the current speaker. This would allow the wearer of the system to

identify and observe the current speaker through the display (viewfinder) of the wearable device.

Butnaru et al. and Potts et al. do not teach

- a video server coupled to the one or more video cameras and operative to: (i) analyze the captured video images to determine which individual has one or more facial features indicative of speech; and (ii) identify the location of the individual who is currently speaking during the event;
- displaying a first visual indicator to the user, in accordance with the display system, in association with the individual identified as the current speaker when the individual is within the field of view of the user;
- displaying a second visual indicator to the user, in accordance with the display system, when the individual identified as the current speaker is not within the field of view of the user.

Van Schyndel teaches capturing images of the users and identifying current speaker based on his/her head and mouth movements and determining the location of that person (Col. 7, lines 35-46). Because video server is inherently just a computer with specialized software, Van Schyndel's computer-based system also reads on this part of the claim.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the wearable system of Butnaru et al. and Potts et al. as taught by Van Schyndel, in order to improve the identification of the speaker, because detection of mouth movements would produce more reliable results at the expense of heavier image processing.

Butnaru et al., Potts et al. and Van Schyndel do not teach:

- displaying a first visual indicator to the user, in accordance with the display system, in association with the individual identified as the current speaker when the individual is within the field of view of the user;
- displaying a second visual indicator to the user, in accordance with the display system, when the individual identified as the current speaker is not within the field of view of the user.

Jhabvala et al. teaches the use of visual indicators to instruct the deaf users about the direction of incoming sounds (Col. 3, lines 2-8).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify a wearable device of Butnaru, modified with a speaker-locating capability taught by Potts et al. and Van Schyndel to utilize visual indicators as taught by Jhabvala et al., in order to allow the wearable system to identify the direction of incoming speech, and if the speaker was not in view, using visual indicators to instruct the user to turn his/her head in the appropriate direction. Similar to the many well-known uses of various indicators in normal life (turn signals in cars, green/red arrows on the intersection lights, highlighting of selected items on Windows, etc), the use of visual indicators with the display would assist deaf people in determining the direction of incoming speech and focusing on the current speaker.

As per claims 25-26, Butnaru et al. disclose a system capable of displaying a variety of visual indicators on the computer screen (FIG.2)

Butnaru et al. do not disclose "visual indicator [that] comprises a change in at least one attribute associated with a representation of the individual identified as the current speaker on the display system", where the attribute is one of color and brightness.

The examiner takes official notice that computer screens have an inherent ability to display various images of varying colors and brightness. In addition, the technique of highlighting objects on the screen to bring them into the focus is well-known to the practitioners in computer arts. For example, Windows operating system changes the color and brightness of file folder images when user selects them.

It would have been obvious to one of ordinary skill in the art at the time the invention was made that display system disclosed by Butnaru et al. to change the color and brightness of the current speaker's image in order to pinpoint the speaker to the user, because the change in the visual representation of the current speaker would quickly get attention of the user and allow him to focus on the speaker.

As per claims 27-28, Butnaru et al. discloses the system that is capable of recognizing the speech using speech recognizer (elem. 55, FIG. 3) and displaying the text content to the user via a display system (Col. 5, lines 49-57).

8. Claim 24 is rejected under 35 U.S.C. 103(a) as being unpatentable over Butnaru et al., Potts et al. and Jhabvala et al., as applied to claim 23, and further in view of Hein et al.

Butnaru et al. do not disclose that a "first visual indicator comprises a marker displayed in proximity to a representation of the individual identified as the current speaker on the display system."

Hein et al. teaches using a colored frame around the speaker's image in order to identify him to the viewer (Col. 7, lines 9-11)

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify visual indicators described by Butnaru et al. as taught by Hein et al. in order to identify the current speaker located in the field of view of the user, because the change in the visual representation of the current speaker would quickly get attention of the user and allow him to focus on the speaker.

Conclusion

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dmitry Brant whose telephone number is (703) 305-8954. The examiner can normally be reached on Mon. - Fri. (8:30am - 5pm).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Ivars Smits can be reached on (703) 306-3011. The fax phone number for the organization where this application or proceeding is assigned to (703) 872-9306. Any inquiry of a general nature or relating to the status of this application or

proceeding should be directed to Tech Center 2600 receptionist whose telephone number is (703) 305- 4700.

DB

2/13/04

TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER